INFORMATION AND CONTROL 37, 182–196 (1978)

Forêts Algébriques et Homomorphismes Inverses

ANDRÉ ARNOLD ET MAX DAUCHET

Université de Lille I, UER IEEA Informatique, BP 36, 59650 Villeneuve d'Ascq, France

Sauf certains cas particuliers, la classe des forêts algébrique n'est pas fermée par homomorphisme linéaire inverse.

INTRODUCTION

Les grammaires algébriques d'arbres et les forêts algébriques définies par Rounds en 1970 (Rounds, 1970a, b) généralisent au cas des algèbres générales libres les notions de grammaires et langages algébriques sur le monoïde libre. Nivat (1973) et d'autres auteurs ont mis en évidence l'intérêt de la notion de forêt algébrique pour l'étude des schémas de programme récursifs.

En 1974, Maibaum (1974) remarque que la propriété de fermeture de la classe des forêts algébriques par homomorphisme inverse n'est pas encore démontrée et que si elle était vraie—ce qu'il estime vraisemblable— alors cette classe de forêts aurait des propriétés de fermeture analogues à celles définissant les familles abstraites de langages introduites par Ginsburg et Greibach (1969).

Nous montrons qu'il n'en est rien en construisant une forêt algébrique F et un homomorphisme linéaire ϕ tels que $\phi^{-1}(F)$ n'est pas une forêt algébrique. La classe des forêts algébriques ne peut donc pas se définir comme une "famille abstraite de forêts," par analogie avec les familles abstraites de langages.

Dans le contre-exemple utilisé, l'homomorphisme ϕ est défini sur un alphabet gradué contenant un symbole de degré 3. Nous démontrons ensuite que si on ne considère que des homomorphismes linéaires définis sur des alphabets ne contenant que des symboles de degré au plus égal à 2, leurs inverses conservent l'algébricité. Il en est de même pour les homomorphismes linéaires alphabétiques; mais nous montrons que dans ces deux cas la condition de linéarité est indispensable.

En définissant Alg_n comme la classe des forêts algébriques dont les arbres ne contiennent pas de symboles de degré strictement supérieur à n (Alg_1 peut alors s'identifier à la classe des langages algébriques), nous avons montré par ailleurs (Arnold et Dauchet, 1975) que pour tout n, Alg_n est l'ensemble des images par transduction linéaire complète (ascendante ou descendante) d'une seule forêt algébrique D_n, mais seules les classes Alg_1 et Alg_2 sont fermées par homomorphismes linéaires inverses. La classe Alg_2 nous semble donc devoir

0019-9958/78/0372–0182$02.00/0

avoir des propriétés qui l'apparentent plus à celle des langages qu'aux autres classes Alg_n.

Cet article comprend 5 parties. La première partie rappelle quelques propriétés élémentaires des grammaires algébriques et des dérivations dans ces grammaires. Dans la seconde on définit la forêt algébrique F, l'homomorphisme ϕ et on donne quelques propriétés de $\phi^{-1}(F)$. La troisième partie est consacrée à la démonstration de la non-algébricité de $\phi^{-1}(F)$. La quatrième partie étudie des homomorphismes inverses particuliers et dans la cinquième partie nous tirons quelques conclusions des résultats obtenus.

La lecture de cet article suppose une connaissance préalable des notions élémentaires de la théorie des langages d'arbres. L'exposé complet de ces notions tiendrait une place à notre avis disproportionnée. Nour préférons renvoyer le lecteur aux articles de Rounds, au cours polycopié d'Engelfriet (1975) et à la thèse de Dauchet (1975) pour les notions générales. En ce qui concerne plus particulièrement les grammaires et forêts algébriques les notions fondamentales se trouvent dans Engelfriet et Schmidt (1975), Maibaum (1974), Rounds (1970a, b).

1. Préliminaires

Un *alphabet gradué* est un couple (Σ, d) où Σ est un ensemble fini et d une application de Σ dans $\mathbb{N}$ appelée le degré. Un alphabet gradué sera le plus souvent noté simplement Σ. Si Σ est un alphabet gradué, pour tout entier n, on note Σ_n l'ensemble $\{\sigma \in \Sigma \mid d(\sigma) = n\}$ des symboles de degré n.

L'ensemble des *arbres* sur un alphabet gradué Σ, noté T_Σ, est défini inductivement par:

$$\Sigma_0 \subset T_\Sigma,$$
$$\text{si } \sigma \in \Sigma_n,\ t_1,\ldots, t_n \in T_\Sigma \text{ alors } \sigma(t_1,\ldots, t_n) \in T_\Sigma.$$

Si E est un ensemble quelconque, l'ensemble des *arbres indexés par E*, noté $T_\Sigma(E)$ est l'ensemble des arbres sur l'alphabet Σ' défini par $\Sigma'_0 = \Sigma_0 \cup E$ et $\forall i > 0$, $\Sigma'_i = \Sigma_i$.

On dira qu'un arbre est *indexé* s'il est indexé par l'ensemble de *variables* $X = \{x_1,\ldots, x_n,\ldots\}$. En particulier, tout arbre de T_Σ est un arbre indexé.

On dira que le symbole f de Σ est la *racine* de t si $t = f(t_1,\ldots, t_n)$.

On fera apparaitre tout les variables figurant dans un arbre t de $T_\Sigma(X)$ en l'écrivant $t(x_{i_1},\ldots, x_{i_p})$. L'arbre noté $t(t_1,\ldots, t_p)$ sera alors obtenu en remplaçant dans $t(x_{i_1},\ldots, x_{i_p})$ la ième occurrence d'une variable par l'arbre (indexé) t_i:

La *profondeur* d'un arbre t, notée $|t|$, est la longueur de sa plus grande branche, définie inductivement par

$$\begin{array}{ll} \text{si } t = x_i \in X, & |t| = 0, \\ \text{si } t = f \in \Sigma_0, & |t| = 1, \\ \text{si } t = f(t_1,\ldots, t_n), & |t| = 1 + \sup(|t_1|,\ldots, |t_n|). \end{array}$$

Un arbre *équilibré* est un arbre dont toutes les branches ont même longueur.

Soit Δ un nouvel alphabet gradué. A toute application φ de Σ dans $T_\Delta(X)$ qui vérifie:

$\forall n \geqslant 0$, $\forall f \in \Sigma_n$, $\phi(f) \in T_\Delta(\{x_1, ..., x_n\})$, on associe une application notée également φ, de $T_\Sigma(X)$ dans $T_\Delta(X)$ appelée *homomorphisme d'arbre*, définie par

$$\forall x_i \in X, \qquad \phi(x_i) = x_i,$$
$$\forall n \geqslant 0, \qquad \forall f \in \Sigma_n, \qquad \forall t_1, ..., t_n \in T_\Sigma(X),$$

$\phi(f(t_1, ..., t_n))$ est l'arbre indexé obtenu en remplaçant chaque variable par $\phi(t_i)$ dans $\phi(f)$.

Remarquons que cette définition est cohérente avec l'identification classique du symbole f de degré n à l'arbre indexé $f(x_1, ..., x_n)$.

EXEMPLE 1. Soient $\Sigma = \{c, a, \#\}$ avec $d(c) = 3$, $d(a) = 1$, et $d(\#) = 0$, $\Delta = \Sigma$.

Soit ϕ défini par

$$\phi(c) = c(a(x_1), x_1, \#),$$
$$\phi(a) = c(x_1, \#, \#),$$
$$\phi(\#) = a(\#).$$

Alors

$$\phi(a(\#)) = b(a(\#), \#, \#)$$

et

$$\phi(c(a(\#), x_1, \#)) = c(a(b(a(\#), \#, \#)), b(a(\#), \#, \#), \#).$$

Un homorphisme ϕ sera dit *linéaire* (resp.: *complet*) si $\forall n > 0$, $\forall f \in \Sigma_n$, toute variable x_i de $\{x_1, ..., x_n\}$ figure au plus une fois (resp.: au moins une fois) dans $\phi(f)$. Il sera dit *strict* si $\forall f \in \Sigma, \phi(f)$ n'est pas un arbre réduit à une variable.

Ainsi l'homomorphisme défini dans l'exemple précédent n'est ni linéaire ni complet mais il est strict. L'homomorphisme défini ci-dessous, et qu'on retrouvera par la suite est linéaire, complet et strict.

EXEMPLE 2. Soient $\Sigma = \{c, b, a, \#\}$ et $\Sigma' = \{b, b_1, b_2, a, \#\}$ avec $d(c) = 3$, $d(b) = d(b_1) = d(b_2) = 2$, $d(a) = 1$, et $d(\#) = 0$.

Soit ϕ défini par

$$\phi(c) = b_1(b_2(x_1, x_2), x_3)),$$
$$\phi(b) = b(x_1, x_2),$$
$$\phi(a) = a(x_1),$$
$$\phi(\#) = \#.$$

Alors

$$\phi(c(a^i(\#), a^i(\#), a^n(\#))) = b_1(b_2(a^i(\#)), a^i(\#)), a^n(\#))$$

où

$$a^i(\#) \text{ est une abbréviation pour } \underbrace{a(a(a \cdots a(\#)\cdots)))}_{i \text{ fois}}.$$

On appelle *grammaire algébrique* (ou "context-free") tout quadruplet $\langle V, \Sigma, S_0, \mathscr{R}\rangle$ où V (ensemble des symboles non terminaux) et Σ (ensemble des symboles terminaux) sont des alphabets gradués; S_0, l'axiome, est un symbole de V de degré 0 et $\mathscr{R}$ est un ensemble fini de règles.

Une règle est un couple $\langle A(x_1, \dots, x_n), t(x_{i_1}, \dots, x_{i_p})\rangle$ écrit $A(x_1, \dots, x_n) \to t(x_{i_1}, \dots, x_{i_p})$ où A est un symbole non terminal de degré n et ou t est un arbre de $T_{\Sigma\cup V}(\{x_1, \dots, x_n\})$.

Soit G une grammaire algébrique $\langle V, \Sigma, S_0, \mathscr{R}\rangle$. On dira que *l'arbre indexé t' de $T_{\Sigma\cup V}(X)$ se dérive immédiatement en l'arbre indexé t″ par application de la règle* $r = A(x_1, \dots, x_n) \to t(x_{i_1}, \dots, x_{i_p})$, ce qui se notera $t' \Rightarrow_r t''$, si

$$t' = t_0(t_1, \dots, t_j, A(u_1, \dots, u_n), t_{j+1}, \dots, t_k),$$

et

$$t'' = t_0(t_1, \dots, t_j, t(u_{i_1}, \dots, u_{i_p}), t_{j+1}, \dots, t_k).$$

On appelle *dérivation de longueur n* toute séquence $t_0 \Rightarrow_{r_1} t_1 \Rightarrow_{r_2} t_2 \cdots \Rightarrow_{r_n} t_n$ telle que pour tout $i \in \{1, \dots, n\}$ $t_{i-1} \Rightarrow_{r_i} t_i$ est une dérivation immédiate.

On dira que *t′ se dérive en t″*, ce qu'on notera $t' \overset{*}{\Rightarrow}_G t''$, s'il existe une dérivation $t_0 \Rightarrow_{r_1} t_1 \cdots \Rightarrow_{r_n} t_n$ telle que $t_0 = t'$ et $t_n = t''$.

On appelle *forêt engendrée par la grammaire G*, notée $F(G)$ l'ensemble $\{t \in T_\Sigma / S_0 \overset{*}{\Rightarrow}_G t\}$.

On sait (Engelfriet et Schmidt, 1975; Rounds, 1969) que si $t' \overset{*}{\Rightarrow}_G t''$, alors t' se dérive en t'' par une dérivation *descendante* ("OI-dérivation" dans Engelfriet et Schmidt), i.e., on ne dérive que des symboles non terminaux qui ne sont pas "précédés" d'un autre symbole non terminal. Aussi à partir de maintenant, on ne considère plus que des dérivations descendantes.

Soit $G = \langle V, \Sigma, S_0, \mathscr{R}\rangle$ une grammaire algébrique. On dira que cette grammaire est sous *forme normale* (Maibaum, 1974) si toutes ses règles sont de la forme:

(1) $A(x_1, \dots, x_n) \to A'(A_1(x_1, \dots, x_n), \dots, A_p(x_1, \dots, x_n))$ (règles croissantes);

(2) $A(x_1, \dots, x_n) \to x_i$ (règles effaçantes);

(3) $A(x_1, \dots, x_n) \to f(x_{i_1}, \dots, x_{i_p})$ où $f \in \Sigma_p$ (règles terminales).

On dira qu'une grammaire algébrique est *totale* si $\forall n$, $\forall A \in V$ de degré n, $\exists t \in T_\Sigma(x_1, \dots, x_n)$ tel que $A(x_1, \dots, x_n) \overset{*}{\Rightarrow} t$; autrement dit tout symbole non terminal peut se dériver en un arbre indexé sur l'alphabet terminal.

Lemme 1.1. *Quelle que soit la forêt algébrique F, il existe une grammaire G totale sous forme normale qui engendre F.*

L'existence d'une grammaire sous forme normale est connue (Maibaum, 1974). Le fait qu'on peut la supposer totale est démontré dans Arnold et Dauchet (1976).

Lemme 1.2. *Quelle que soit la grammaire algébrique $G = \langle V, \Sigma, S_0, \mathscr{R}\rangle$, il existe un entier p tel que $\forall n$, $\forall A \in V$ de degré n, si il existe un arbre t de $T_\Sigma(\{x_1,..., x_n\})$ de profondeur strictement supérieure à p tel que $A(x_1,..., x_n) \overset{*}{\Rightarrow} t$, alors l'ensemble $\{t \in T_\Sigma(x_1,..., x_n) \mid A(x_1,..., x_n) \overset{*}{\Rightarrow} t\}$ est un ensemble infini.*

Preuve. A chaque symbole non terminal A de V on associe l'ensemble $F(G, A) = \{t \in T_\Sigma(x_1,..., x_n) \mid A(x_1,..., x_n) \overset{*}{\Rightarrow} t$. On pose $V' = \{A \in V \mid \mathrm{Card}(F(G, A)) < +\infty\}$. L'ensemble $F_0 = \bigcup_{A\in V'} F(G, A)$ est alors un ensemble fini. Il suffit de prendre $p = \max\{\mid t \mid t \in F_0\}$. Il est immédiat que si $(A(x_1,..., x_n) \overset{*}{\Rightarrow} t$ avec $\mid t \mid > p$ alors $A \notin V'$ et $F(G, A)$ est un ensemble infini. C.Q.F.D.

Soit $G = \langle V, \Sigma, S_0, \mathscr{R}\rangle$ une grammaire algébrique et soit $\delta: S_0 \Rightarrow v_1 \Rightarrow v_2 \cdots \Rightarrow v_n$, où $v_n \in T_\Sigma$, une dérivation de v_n dans G. La sous-dérivation $\delta': S_0 \Rightarrow v_1 \cdots \Rightarrow v_i$ avec $i < n$ sera dite *initiale* si la racine de chacun des v_j pour $j < i$ est étiquetée par un symbole non terminal, et dans ce cas v_{j+1} est obtenu en dérivant ce symbole non terminal.

Lemme 1.3. *Quelle que soit la grammaire algébrique totale G sous forme normale, il existe une grammaire G' totale sous forme normale qui engendre la même forêt F que G et telle que $\forall t \in F$, il existe dans G' une dérivation $\delta: S_0 \Rightarrow v_1 \cdots \Rightarrow t$ telle que dans toute sous dérivation initiale $\delta'_i: S_0 \Rightarrow v_1 \cdots \Rightarrow v_i$ on passe de v_j à v_{j+1} en appliquant à la racine de v_j une règle croissante (i.e., de la forme $A(x_1,..., x_n) \to A'(A_1(x_1,..., x_n),..., A_p(x_1,..., x_n))$.*

Preuve. Nous avons déjà montré (Arnold et Dauchet, 1976) qu'on pouvait ne pas utiliser de règles effaçantes dans une dérivation initiale à condition de rajouter à la grammaire G des règles de la forme $A(x_1,..., x_n) \to A'(x_1,..., x_n)$. Par un procédé analogue on montre qu'on peut toujours se passer de telles règles à condition d'ajouter à G de nouvelles règles croissantes, effaçantes ou finales. Comme cetter transformation n'introduit pas de nouveaux symboles non terminaux, la grammaire ainsi obtenue reste totale. C.Q.F.D.

Lemme 1.4. *Soit G une grammaire algébrique qui engendre la forêt F. Quel que soit $t \in F$, il existe une dérivation $\delta: S_0 \Rightarrow v_1 \Rightarrow v_2 \cdots \Rightarrow t$ telle que pour toute sous-dérivation initiale $\delta': S_0 \Rightarrow v_1 \cdots \Rightarrow v_n$ de longueur n, v_n est un arbre équilibré de profondeur $n + 1$.*

Ce résultat est une conséquence immédiate du Lemme 1.3.

Soit δ: $S_0 \to v_1 \cdots \Rightarrow t$ une dérivation. On appelle sous-dérivation initiale *maximale* la sous-dérivation initiale δ_m: $S_0 \Rightarrow v_1 \cdots \Rightarrow v_m$ de δ telle que v_{m+1} est obtenue en dérivant la racine de v_m par une règle finale (i.e., $A(x_1, ..., x_n) \to f(x_{i_1}, ..., x_{i_p})$. Il est clair que la sous-dérivation initiale maximale d'une dérivation existe et est unique (elle peut être éventuellement de longueur 0 si δ: $S_0 \Rightarrow t \in T_\Sigma$) (voir Arnold et Dauchet, 1976).

De plus toute sous-dérivation initiale de δ est une sous-dérivation initiale de cette sous-dérivation initiale maximale.

LEMME 1.5. *Soit* δ_m: $S_0 \Rightarrow v_1 \Rightarrow v_m$ *la sous-dérivation initiale maximale de la dérivation* δ: $S_0 \Rightarrow v_1 \cdots \Rightarrow v_m \Rightarrow t$, *ou* δ *vérifie les conclusions du Lemme* 1.3. *Alors à tout entier* $k < m$ *on peut associer un arbre indexé* $t_k = t_k(x_{j_1}, ..., x_{j_p})$ *qui vérifie*:

—*il existe des arbres indexés* $t'_1, ..., t'_p$ *tels que* $t = t_k(t'_1, ..., t'_p)$,

—$A_k(x_1, ..., x_n) \overset{*}{\Rightarrow} t_k(x_{j_1}, ..., x_{j_p})$ *où* A_k *est la racine de* v_k.

De plus si $k' < k$ *alors il existe* $t''_1, ..., t''_p$ *tels que* $t_{k'} = t_k(t''_1, ..., t''_p)$.

Preuve. On démontre aisément par récurrence sur la longueur des dérivations que si $u(u_1, ..., u_n) \overset{*}{\Rightarrow} t$, alors $t = t_0(t_1, ..., t_m)$ avec $u(x_1, ..., x_n) \overset{*}{\Rightarrow} t_0(x_{i_1}, ..., x_{i_m})$ et $\forall j \leqslant m$, $u_{i_j} \overset{*}{\Rightarrow} t_j$. Il suffit alors d'appliquer ce résultat à la dérivation considérée.

Pour démontrer la seconde partie du lemme on tient également compte du fait que la dérivation $S \Rightarrow v_1 \Rightarrow \cdots \Rightarrow v_{k'} \cdots \Rightarrow v_k$ est initiale. C.Q.F.D.

2. DÉFINITION DU CONTRE-EXEMPLE

On considère l'alphabet $\Sigma' = \{b, b_1, b_2, a, \#\}$ de l'exemple 2 et la grammaire algébrique $G = \langle V', \Sigma', S'_0, \mathscr{R}\rangle$ où $V' = \{S'_0, S, A, D\}$ et dont les règles sont

$$S'_0 \to A(b_2(\#, \#), a(\#));$$

$$A(x, y) \to A(S(x, y), a(y)); \qquad A(x, y) \to D(b_1(x, y));$$

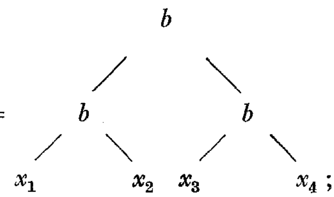
$$D(x) \to D(b(x, x)); \qquad D(x) \to b(x, x);$$

$$S(x, y) \to x; \qquad S(x, y) \to b_2(y, y).$$

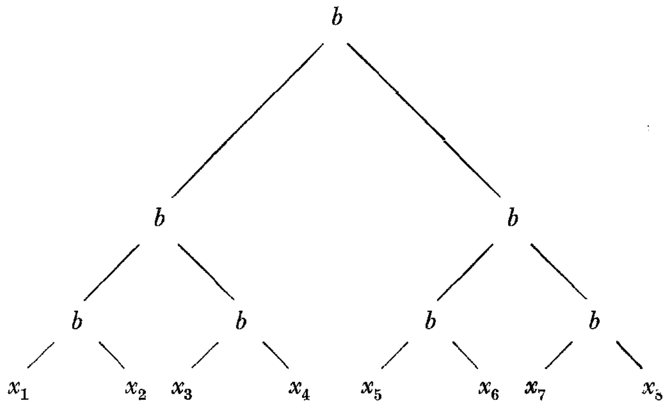
Pour $m > 1$, on notera $B_m = B_m(x_1, x_2, ..., x_{2^m})$ l'arbre équilibré de profondeur m de $T_b(X)$ dont les 2^m feuilles sont étiquetées de gauche à droite par $x_1, ..., x_{2^m}$.

EXEMPLE.

$B_2 =$



$B_3 =$



Pour $n > 1$ on note E'_n l'ensemble d'arbres $\{b_1(b_2(a^i\#, a^i\#), a^n\#) \mid i < n\}$ et pour $n \geqslant 1$, $m \geqslant 1$ on note $F'_{m,n}$ l'ensemble d'arbres $\{B_m(t_1, ..., t_{2^m}) \mid \forall j \in \{1, ..., 2^m\}, t_j \in E'_n\}$.

Il est facile de vérifier que la forêt engendrée par G est $F' = \bigcup_{m \geqslant 1, n \geqslant 1} F'_{m,n}$.

On considère maintenant l'alphabet $\Sigma = \{c, b, a, \#\}$ de l'exemple 2 et l'homomorphisme ϕ, linéaire, complet et strict, de $T_\Sigma(X)$ dans $T_{\Sigma'}(X)$ défini dans ce même exemple.

En posant $E_n = \{c(a^i\#, a^i\#, a^n\#) \mid i < n\}$ pour $n \geqslant 1$ et $F_{m,n} = \{B_m(t_1, ..., t_{2^m}) \mid \forall j \in \{1, ..., 2^m\}, t_j \in E_n\}$ pour $m \geqslant 1$ et $n \geqslant 1$, il est clair que

$$\phi^{-1}(F') = F = \bigcup_{\substack{m \geqslant 1 \\ n \geqslant 1}} F_{m,n}.$$

La forêt F ainsi définie vérifie les deux propriétés suivantes:

PROPOSITION 2.1. *Soient $t \in F$, $t_1(x) \in T_\Sigma(X)$, qui ne contient qu'une seule occurrence de x, et k un entier positif ou nul tels que $t_1(a^k\#) = t$. Alors $\{t' \mid t_1(t') \in F\} = \{a^k\#\}$.*

Preuve. Puisque $t_1(a^k\#) \in F$ alors $t_1(x)$ s'écrit

(1) $t'_1(c(a^i\#, a^i\#, a^{k'}(x))$ avec $k + k' = n$ ou

(2) $t'_1(c(a^{i'}(x), a^i\#, a^n\#))$ avec $i' + k = i$ ou

(3) $t'_1(c(a^i\#, a^{i'}(x), a^n(\#))$ avec $i' + k = i$.

Si $t_1(a^{k''}\#) \in F$ alors dans les cas 2 et 3 on a nécessairement $i' + k'' = i$ et donc $k = k''$, et dans le cas 1, $k' + k'' = n$ et donc encore $k = k''$. C.Q.F.D.

PROPOSITION 2.2. *Soient $t \in F$, $t_1(x) \in T_\Sigma(X)$ non réduit à la variable x et ne contenant qu'une seule occurrence de x et $t_2 \in T_\Sigma$ tels que $t = t_1(t_2)$. Alors l'ensemble $\{t' \mid t_1(t') \in F\}$ est fini.*

Preuve. Si $t_2 \in a^*\#$ on est ramené à la proposition précédente. Sinon il existe m et n tels que $t \in F_{m,n}$, et il existe $m' < m$ tel que $t_2 \in F_{m',n}$. D'après la définition de F, si $t_1(t') \in F$ alors $t' \in F_{m',n}$ et $F_{m',n}$ est bien un ensemble fini. C.Q.F.D.

3. LA FORÊT F N'EST PAS ALGÉBRIQUE

Supposons que F soit une forêt algébrique.

Soit $G = \langle V, \Sigma, S_0, \mathscr{R}\rangle$ une grammaire algébrique totale sous forme normale qui vérifie la conclusion du Lemme 1.3 et qui engendre F. Soit p l'entier défini dans le Lemme 1.1, soit $m = \operatorname{Card}(V)$ le nombre de symboles non-terminaux de G et soit d le degré maximum de V (i.e., $\max\{d(A) \mid A \in V\}$).

PROPOSITION 3.1. *Soit $t \in F_{k,n} \subset F$ et soit δ: $S_0 \Rightarrow v_1 \Rightarrow v_2 \cdots \Rightarrow t$ une dérivation de t dans G. Soit δ': $S_0 \Rightarrow v_1 \cdots \Rightarrow v_{i_0} = A(v'_1, \ldots, v'_n)$ une sous-dérivation initiale de δ de longueur i_0 telle que:*

(i) $t = t_0(t_1, \ldots, t_q)$,

(ii) $\forall j \in \{1, \ldots, q\}$, $\quad t_j \in a^*\#$ *et* $v'_{i_j} \overset{*}{\Rightarrow} t_j$,

(iii) $A(x_1, \ldots, x_n) \overset{*}{\Rightarrow} t_0(x_i, \ldots, x_{i_q})$,

alors $\forall j \in \{1, \ldots, q\}$

(α) $\{t' \in T_\Sigma \mid v'_{i_j} \overset{*}{\Rightarrow} t'\} = \{t_j\}$,

(β) $|t_j| \leqslant p \times i_0$.

Preuve. Comme δ' est une sous-dérivation initiale de δ, $S_0 \overset{*}{\Rightarrow} A(v'_1, \ldots, v'_n) \overset{*}{\Rightarrow} t_0(v'_{i_1}, \ldots, v'_{i_q}) \overset{*}{\Rightarrow} t_0(t_1, \ldots, t_q) = t$.

Soit $j \in \{1, \ldots, q\}$. Alors $v'_{i_j} \overset{*}{\Rightarrow} t_j$. Si $v'_{i_j} \overset{*}{\Rightarrow} t'$ alors $S_0 \overset{*}{\Rightarrow} t_0(t_1, \ldots, t_{j-1}, t', t_{j+1}, \ldots, t_q)$. Cet arbre appartient à F et en appliquant la proposition 2.1 à $t_0(t_1, \ldots, t_{j-1}, x, t_{j+1}, \ldots, t_q)$, $t' = t_j$.

Si $|t_j| > p \times i_0$, comme d'après le Lemme 1.4 v'_i est un arbre équilibré de profondeur i_0, la dérivation $v'_{i_j} \overset{*}{\Rightarrow} t_j$ s'écrit nécessairement

$$v'_{i_j} \overset{*}{\Rightarrow} a^k(A(v''_1, \ldots, v''_m)) \overset{*}{\Rightarrow} a^k(a^{k'}(v''_i)) \overset{*}{\Rightarrow} t_j,$$

avec $A(x_1, \ldots, x_m) \overset{*}{\Rightarrow} a^{k'}(x_i)$, et $k' > p$. D'après le Lemme 1.2 on peut dériver une infinité d'arbres à partir de $A(x_1, ..., x_m)$ et on peut donc dériver de v'_i une infinité d'arbres, puisque la grammaire G est totale, ce qui contredit α).

C.Q.F.D.

Soit maintenant un entier k tel que $2^k > p.m.d.$ et posons $r = 2^k$. On considère l'arbre $t = B_k(\hat{t}_1, ..., \hat{t}_r)$ où pour tout $j \in \{1, ..., r\}$ $\hat{t}_j = c(a^{j-1}\#, a^{j-1}\#, a^r\#)$. Cet arbre appartient bien à $F_{k,r} \subseteq F$.

(A) Soit $\delta: S_0 \Rightarrow v_1 \Rightarrow v_2 \cdots \Rightarrow t$ une dérivation de t dans G et soit $\delta': S_0 \Rightarrow v_1 \cdots \Rightarrow v_{i_0}$ la sous-dérivation initiale maximale de δ; en vertu du Lemme 1.5 on associe à chaque $i \leqslant i_0$ un arbre indexé t_i. D'après le Lemme 1.5 et comme $t_{i_0} = b(x_{i_1}, x_{i_2})$, il existe donc un plus grand entier j inférieur à i_0 tel que t_j est de la forme

$$B_k(c(t'_1(x_{i_1}), t'_2(x_{i_2}), t'_3(x_{i_3})), c(t'_4(x_{i_4}), t'_5(x_{i_5}), t'_6(x_{i_6})),\\ ..., c(t'_{3r-2}(x_{i_{3r-2}}), t'_{3r-1}(x_{i_{3r-1}}), t'_{3r}(x_{i_{3r}}))).$$

On pose $l(t_j) = \inf_{1 \leqslant q \leqslant r} | t'_{3q} |$.

PROPOSITION 3.2.

$$l(t_j) < p.$$

Preuve. Si $v_j = A(u_1, ..., u_n)$ alors la dérivation δ s'écrit $S_0 \overset{*}{\Rightarrow} A(u_1, ..., u_n) \Rightarrow D(D_1(u_1, ..., u_n), ..., D_m(u_1, ..., u_n)) \overset{*}{\Rightarrow} t$ avec $A(x_1, ..., x_n) \Rightarrow D(D_1(x_1, ..., x_n), ..., D_m(x_1, ..., x_n)) \overset{*}{\Rightarrow} t_j$ et où $D(x_1, ..., x_m) \overset{*}{\Rightarrow} t_{j+1}$; l'arbre t_{j+1} n'ayant pas la forme indiquée ci dessus.

Il existe donc $t'' = t'(c(t'_{3i+1}(x_{s_1}), t'_{3i+2}(x_{s_2}), t'_{3i+3}(x_{s_3}))$ et $k \in \{1, ..., m\}$ et tel que $D_k(x_1, ..., x_n) \overset{*}{\Rightarrow} t''$.

Mais $| t'' | > | t'_{3i+3} | \geqslant l(t_j)$. Si $l(t_j) \geqslant p$, alors de D_k on peut dériver une infinité d'arbres, et comme la grammaire G est totale on pourra aussi en dériver une infinité à partir de $D_k(u_1, ..., u_n)$; ce qui est impossible d'après la proposition 2.2 appliquée à $\hat{t}(x)$ tel que $\hat{t}(t'') = t$. C.Q.F.D.

(B) Supposons que la sous-dérivation initiale $\delta'': S_0 \Rightarrow v_1 \Rightarrow v_1 \cdots \Rightarrow v_j = A(u_1, ..., u_n)$ extraite de δ', où δ' et j sont définis dans la partie A ci-dessus, soit de longueur supérieure à m.

Il existe donc i et i' tels que les racines de v_i et $v_{i'}$ soient étiquetées par le même symbole non terminal D. La dérivation δ'' s'écrit donc $S_0 \overset{*}{\Rightarrow} D(v'_1, ..., v_s') \overset{*}{\Rightarrow} D(v''_1, ..., v''_s) \overset{*}{\Rightarrow} A(u_1, ..., u_n)$.

Comme cette dérivation est initiale, il existe donc une dérivation initiale $D(x_1, ..., x_s) \overset{*}{\Rightarrow} A(u'_1, ..., u'_n)$ de même longueur que la dérivation $D(v''_1, ..., v''_s) \overset{*}{\Rightarrow} A(u_1, ..., u_n)$, telle que u_k est obtenu en remplaçant dans u'_k chaque variable x_i par l'arbre v''_i. On peut alors obtenir une dérivation initiale, encore de même longueur, $D(v'_1, ..., v'_s) \overset{*}{\Rightarrow} A(u''_1, ..., u''_n)$, où u''_k est obtenu en remplaçant dans u'_k la variable x_i par v'_i.

La dérivation $S_0 \overset{*}{\Rightarrow} D(v'_1, \ldots, v'_s) \overset{*}{\Rightarrow} A(u''_1, \ldots, u''_n)$ est encore une dérivation initiale, de longueur strictement inférieure à celle de δ''.

En réitérant ce procédé, on peut fabriquer une dérivation $S_0 \overset{*}{\Rightarrow} A(w_1, \ldots, w_n)$ de longueur strictement inférieure à m. Mais alors, d'après la définition de t_j, $S_0 \overset{*}{\Rightarrow} A(w_1, \ldots, w_n) \overset{*}{\Rightarrow} t_j(w_{i_1}, \ldots, w_{i_{3r}})$ et, comme la grammaire est totale, $\forall q \in \{1, \ldots, 3r\}$, il existe t''_q tel que $w_{i_q} \overset{*}{\Rightarrow} t''_q$. Il s'ensuit que $t_j(t''_1, \ldots, t''_{3r}) \in F$. Donc pour un élément q quelconque de $\{1, \ldots, 3r\}$, $t''_q \in a^*\#$. De plus la dérivation $S_0 \overset{*}{\Rightarrow} A(w_1, \ldots, w_n)$ étant de longueur strictement inférieure à m, w_{i_q} est un arbre équilibré de profondeur strictement inférieure à m. D'après la Proposition 3.1. $|t''_q| \leqslant (m-1)p$.

(C) Comme

$$t_j = B_k(c(t'_1(x_{i_1}), t'_2(x_{i_2}), t'_3(x_{i_3})), \ldots, c(t'_{3r-2}(x_{i_{3r-2}}), t'_{3r-1}(x_{i_{3r-1}}), t'_{3r}(x_{i_{3r}}))),$$

$$B_k(c(t'_1(t''_1), t'_2(t''_2), t'_3(t''_3)), \ldots, c(t'_{3r-2}(t''_{3r-2}), t'_{3r-1}(t''_{3r-1}), t'_{3r}(t''_{3r})))$$

est un élément de F.

D'après la définition de F on a donc

$$\begin{aligned} |t'_3(t''_3)| &= |t'_3| + |t''_3| = |t'_6(t''_6)| = |t'_6| + |t''_6| = \cdots \\ &= |t'_{3r}(t''_{3r})| = |t'_{3r}| + |t''_{3r}|. \end{aligned}$$

On posera cette quantité égale à n'. Or $l(t_j) < p$ et d'après la définition de $l(t_j)$ il existe $q \in \{1, \ldots, r\}$ tel que $l(t_j) = |t'_{3q}|$. Comme $|t''_{3q}| \leqslant (m-1)p$, on en déduit que $n' = |t'_{3q}| + |t''_{3q}| \leqslant (m-1)p + l(t_j) < mp$. Par ailleurs et toujours d'après la définition de F, pour tout $q \in \{1, \ldots, 3r\}$, $|t'_q(t''_q)| = n' < mp$ ce qui entraine que $|t'_q| < mp$. On en déduit que l'ensemble $\{|t'_q| \mid 1 \leqslant q \leqslant 3r\}$a au plus mp éléments.

(D) Rappelons que la dérivation δ de $t = B_k(\hat{t}_1, \ldots, \hat{t}_r)$ s'écrit $S_0 \overset{*}{\Rightarrow} v_j \overset{*}{\Rightarrow} t_j(u_{j_1}, \ldots, u_{i_{3r}}) \overset{*}{\Rightarrow} t$ avec $v_j = A(u_1, \ldots, u_n)$.

Pour tout $q \in \{1, \ldots, 3r\}$ il existe donc $\bar{t}_q$ tel que $u_{i_q} \overset{*}{\Rightarrow} \bar{t}_q$ et $t = t_j(\bar{t}_1, \ldots, \bar{t}_{3r})$. Comme $\bar{t}_q \in a^*\#$, d'après la proposition 3.1, on ne peut dériver qu'un seul élément de T_Σ à partir de u_{i_q} d'où $\{\bar{t}_q \mid 1 \leqslant q \leqslant 3r\}$ est un ensemble à au plus n éléments. Mais

$$\begin{aligned} t &= t_j(\bar{t}_1, \ldots, \bar{t}_{3r}) \\ &= B_k(c(t'_1(\bar{t}_1), t'_2(\bar{t}_2), t'_3(\bar{t}_3)), \ldots, c(t'_{3r-2}(\bar{t}_{3r-2}), t'_{3r-1}(\bar{t}_{3r-1}), t'_{3r}(\bar{t}_{3r}))) \end{aligned}$$

avec

$$\begin{aligned} t'_3(\bar{t}_3) &= t'_6(\bar{t}_6) \cdots = t'_{3r}(\bar{t}_{3r}) = a^r\#, \\ t'_1(\bar{t}_1) &= t'_2(\bar{t}_2) = \#, \\ t'_4(\bar{t}_4) &= t'_5(\bar{t}_5) = a\#, \\ &\cdots \\ t'_{3r-2}(\bar{t}_{3r-2}) &= (t'_{3r-1}) = a^{r-1}\# \end{aligned}$$

c'est-à-dire que $\mathrm{Card}\{|t_i'(\bar{t}_i)| \mid 1 \leqslant i \leqslant 3r\} = r$. Or, comme $|t'_i(\bar{t}_i)| = |t'_i| + |\bar{t}_i|$,

$$r = \mathrm{Card}\{|t'_i(\bar{t}_i)| \mid 1 \leqslant i \leqslant 3r\} \leqslant \mathrm{Card}\{|t'_i| \mid 1 \leqslant i \leqslant 3r\} \\ \times \mathrm{Card}\{|\bar{t}_i| \mid 1 \leqslant i \leqslant 3r\} \leqslant mp \times n.$$

Mais n étant le degré du symbole non terminal A, par définition $n \leqslant d$ ce qui implique $r \leqslant m.p.d.$ Or on a justement choisi $r > m.p.d.$ d'où une contradiction. La forêt F n'est donc pas algébrique; d'où le théorème:

THÉORÈME 3.1. *La classe des forêts algébriques n'est pas fermée par homomorphisme linéaire inverse.*

On remarquera qu'on a en fait démontré le théorème pour la classe la plus restreinte des homomorphismes linéaires complets stricts dont les propriétés particulières sont étudiées dans la thèse de Dauchet (1975).

4. LE CAS D'HOMOMORPHISMES PARTICULIERS

On appelle *démarquage propre* tout homorphisme π de $T_\Sigma(X)$ dans $T_\Delta(X)$ défini par $\forall n$, $\forall \sigma \in \Sigma_n$, $\pi(\sigma(x_1, ..., x_n)) = \delta(x_1, ..., x_n)$ où $\delta \in \Delta_n$.

On appelle *démarquage* tout homomorphisme ϕ de $T_\Sigma(X)$ dans $T_\Delta(X)$ défini par $\forall n$, $\forall \sigma \in \Sigma_n$, $\phi(\sigma(x_1, ..., x_n)) = \delta(x_{i_1}, ..., x_{i_p})$ où $\delta \in \Delta_p$ ou $\phi(\sigma(x_1, ..., x_n)) = x_i$.

On remarque que les démarquages propres sont des cas particuliers de démarquages linéaires complets stricts.

On appelle *homomorphisme marqué* tout homomorphisme linéaire complet strict ϕ de $T_\Sigma(X)$ dans $T_\Delta(\Sigma)$ qui vérifie:

(1) Il existe une partition $\{\Delta_\sigma\}_{\sigma\in\Sigma}$ de Δ telle que $\forall n$, $\forall \sigma \in \Sigma_n$, $\phi(\sigma(x_1, ..., x_n)) \in T_{\Delta_\sigma}(\{x_1, ..., x_n\})$.

(2) $\forall n$, $\forall \sigma \in \Sigma_n$, tous les noeuds de $\phi(\sigma(x_1, ..., x_n))$ qui ne sont pas étiquetés par des variables ont des étiquettes différentes.

LEMME 4.1. *Tout homomorphisme (linéaire) est le composé d'un démarquage (linéaire), d'un homomorphisme marqué et d'un démarquage propre.*

Preuve. Soit ϕ un homomorphisme de $T_\Sigma(X)$ dans $T_\Delta(X)$. On définit α, $\bar{\phi}$ et π par $\forall n$, $\forall \sigma \in \Sigma_n$

si $\phi(\sigma(x_1, ..., x_n)) = x_i$ alors $\alpha(\sigma(x_1, ..., x_n)) = x_i$,
si $\phi(\sigma(x_1, ..., x_n)) = t(x_{i_1}, ..., x_{i_p})$ alors

—$\alpha(\sigma(x_1, ..., x_n)) = \bar{\sigma}(x_{i_1}, ..., x_{i_p})$,

—à chaque noeud de t non étiqueté par une variable on associe un numéro différent et on forme $\bar{t}$ en remplaçant l'étiquette δ du noeud de numéro i par l'étiquette $\delta_{\sigma,i}$ dans t et on pose $\bar{\phi}(\bar{\sigma}(x_1, ..., x_p)) = \bar{t}(x_1, ..., x_p)$

—$\pi(\delta_{\sigma,i}(x_1, ..., x_q)) = \delta(x_1, ..., x_q)$.

Il est immédiat par construction que α est un démarquage (linéaire si ϕ est linéaire), que $\bar{\phi}$ est un homomorphisme marqué et π un démarquage propre et que $\phi = \pi \circ \bar{\phi} \circ \alpha$. C.Q.F.D.

THÉORÈME 4.1. *La classe des forêts algébriques est fermée par démarquage linéaire inverse.*

Preuve. Soit ϕ un démarquage linéaire de $T_\Sigma(X)$ dans $T_\Delta(X)$. Soit $G = \langle V, \Delta, S_0, \mathscr{R}\rangle$ une grammaire algébrique, qu'on peut supposer sous forme normale, qui engendre la forêt algébrique F de T_Δ.

Soit $G' = \langle V', \Sigma, Z_0, \mathscr{R}'\rangle$ une grammaire régulière engendrant la forêt reconnaissable T_Σ. On considère la grammaire algébrique $G'' = \langle V'', \Sigma, Y_0, \mathscr{R}''\rangle$ définie par

$V'' = V \cup V' \cup \{E\}$ où E est un symbole de degré 1 n'appartenant pas à $V \cup V'$,

$\mathscr{R}'' = \mathscr{R}_c \cup \mathscr{R}' \cup \bar{\mathscr{R}}$ où $\mathscr{R}_c$ est l'ensemble des règles croissantes de $\mathscr{R}$ et où $\bar{\mathscr{R}}$ est défini de la façon suivante:

(1) Si $A(x_1, ..., x_n) \to x_i$ est une règle de $\mathscr{R}_e \subset \mathscr{R}$, alors $A(x_1, ..., x_n) \to E(x_i)$ est une règle de $\bar{\mathscr{R}}$.

(2) $E(x) \to x$ et $E(x) \to E(E(x))$ sont des règles de $\bar{\mathscr{R}}$.

(3) Si $A(x_1, ..., x_n) \to \delta(x_{i_1}, ..., x_{i_p})$ est une règle de $\mathscr{R}_f \subset \mathscr{R}$, alors $A(x_1, ..., x_n) \to E(\sigma(u_1, ..., u_m))$ est une règle de $\bar{\mathscr{R}}$ à condition que

(1) Si $\phi(\sigma(y_1, ..., y_m)) = \delta(y_{j_1}, ..., y_{j_p})$ alors $u_{j_1}, ..., u_{j_p}$ sont des variables et les autres u_k sont égaux à Z_0.

(2) $\phi(\sigma(u_1, ..., u_m)) = \delta(x_{i_1}, ..., x_{i_p})$. Ces deux conditions sont compatibles à cause de la linéarité de ϕ.

EXEMPLE. Si r: $A(x, y, z) \to \delta(x, x, z)$ est une règle de $\mathscr{R}_f$ et si $\phi(\sigma(x_1, x_2, x_3, x_4)) = \delta(x_2, x_3, x_4)$ alors la règle $A(x, y, z) \to \sigma(Z_0, x, x, z)$ est une règle de $\bar{\mathscr{R}}$.

(4) $E(x) \to \sigma(u_1, ..., u_m)$ est une règle de $\bar{\mathscr{R}}$ à condition que

(1) $\phi(\sigma(x_1, ..., x_m)) = x_{i_0}$,

(2) pour $j \in \{1, ..., m\}$,

$$u_j = \begin{cases} x & \text{si } j = i_0, \\ Z_0 & \text{sinon.} \end{cases}$$

La forêt engendrée par cette grammaire est bien $\phi^{-1}(F)$ car on démontre aisément que si $A(x_1, ..., x_n) \overset{*}{\Rightarrow}_G t(x_{i_1}, ..., x_{i_p})$ alors $\forall t' \in T_\Sigma(x_{i_1}, ..., x_{i_p})$ tel que $\phi(t') = t$, $A(x_1, ..., x_n) \overset{*}{\Rightarrow}_{G''} t'$. C.Q.F.D.

On dira qu'un alphabet gradué Σ est *dyadique* si $\forall n > 2$, $\Sigma_n = \varnothing$. On dira qu'un homomorphisme de $T_\Sigma(X)$ dans $T_\Delta(X)$ est *semi-dyadique* (resp. *dyadique*) si Σ (resp. Σ et Δ) sont des alphabets dyadiques.

LEMME 4.2. *Si ϕ est un homomorphisme semi-dyadique marqué de $T_\Sigma(X)$ dans $T_\Delta(X)$, il existe un homomorphisme linéaire ϕ' de $T_\Sigma(X)$ dans $T_\Delta(X)$ tel que $\forall t \in T_\Sigma$, $\phi'(\phi(t)) = t$.*

Preuve. Comme ϕ est marqué, il existe une partition $\{\Delta_\sigma\}_{\sigma\in\Sigma}$ de Δ. Il suffit de définir ϕ_σ' sur Δ_σ pour chaque σ de telle sorte que $\phi_\sigma'(\phi(\sigma(x, y))) = \sigma(x, y)$. ϕ' sera alors la "somme disjointe" des ϕ_σ' et on aura bien $\phi'(\phi(t)) = t$.

(1) Soit $\sigma \in \Sigma_0$; soit δ l'étiquette de la racine de $\phi(\sigma)$; on pose $\phi'(\delta(x_1, ..., x_p)) = \sigma$.

(2) Soit $\sigma \in \Sigma_1$; alors $\phi(\sigma(x)) = t(x)$. Comme t n'est pas réduit à une variable, $t(x)$ peut s'écrire de façon unique sous la forme: $t_0(t_1(\cdots(t_n(x))\cdots))$, n pouvant être nul, où pour tout $j \in \{0,..., n\}$, $t_j(x) = \delta_j(t_1^{(j)},..., t_{k_j-1}^{(j)}, x, t_{k_j+1}^{(j)},..., t_{p_j}^{(j)})$ avec $\delta_j \in \Delta_{p_j}$.

De plus les symboles $\delta_0, ..., \delta_n$ sont tous différents puisque ϕ est marqué. On pose alors

$$\phi'(\delta_0(x_1, ..., x_{p_0})) = \sigma(x_{k_0}),$$

et pour tout $j \in \{1,..., n\}$ $\phi(\delta_j(x_1, ..., x_p)) = x_{k_j}$.

On a bien $\phi'(\phi(\sigma(x)) = \sigma(x)$.

(3) Soit $\sigma \in \Sigma_2$. Alors $\phi(\sigma(x, y)) = t(x, y)$ ou $t(y, x)$. L'arbre $t(x_1, x_2)$ s'écrit de façon unique sous la forme $t_1(t_2(t_3(x_1)), t_4(x_2)))$ où les arbres $t_1(x)$, $t_3(x)$ et $t_4(x)$ peuvent être réduits à des variables et où $t_2(x_1, x_2)$ a la forme $\delta(t'_1, ..., t'_{i-1}, x_1, t'_{i+1}, ..., t'_{j-1}, x_2, t'_{j+1}, ..., t'_p)$ avec $\delta \in \Delta_p$.

On démontre l'existence et l'unicité de cette décomposition par induction sur la profondeur de $t(x_1, x_2)$ en remarquant que $t(x_1, x_2)$ ne peut avoir que l'une des deux formes: $\delta_0(u_1, ..., u_i(x_1, x_2), ..., u_q)$ ou $\delta_0(u_1, ..., u_i(x_1), ..., u_j(x_2), ..., u_q)$. En remarquant une nouvelle fois que tous les noeuds de t ont des étiquettes différentes, on peut définir ϕ' de façon à ce que $\phi'(t_1(x)) = \phi'(t_2(x)) = \phi'(t_3(x)) = x$ et on pose

$$\phi'(\delta(x_1, ..., x_p)) = \begin{cases} \sigma(x_i, x_j) & \text{si } \phi(\sigma(x, y)) = t(x, y), \\ \sigma(x_j, x_i) & \text{si } \phi(\sigma(x, y)) = t(y, x). \end{cases}$$

On aura donc encore $\phi'(\phi(\sigma(x, y)) = \sigma(x, y)$. C.Q.F.D.

THÉORÈME 4.2. *La classe des forêts algébriques est fermée par homomorphisme semi-dyadique linéaire inverse.*

Preuve. Soit ϕ un homomorphisme semi-dyadique linéaire de $T_\Sigma(X)$ dans $T_\Delta(X)$. D'après le Lemme 4.1, il existe deux alphabets Γ et Ω, un démarquage linéaire α de $T_\Sigma(X)$ dans $T_\Gamma(X)$, un homomorphisme marqué $\bar{\phi}$ de T_Γ dans T_Ω et un démarquage propre π de $T_\Omega(X)$ dans $T_\Delta(X)$ tels que $\phi = \pi \circ \bar{\phi} \circ \alpha$. Par construction de α dans le Lemme 4.1, si Σ est un alphabet dyadique, Γ est aussi un alphabet dyadique. Comme la classe des forêts algébriques est fermée par démarquage linéaire inverse, il suffit de montrer que si F est une forêt algébrique de T_Ω, $\bar{\phi}^{-1}(F)$ est une forêt algébrique. D'après le Lemme 4.2, il existe un homomorphisme linéaire ϕ' tel que $\forall t \in T_\Gamma, \phi'(\bar{\phi}(t)) = t$. On en déduit que $\bar{\phi}^{-1}(F) = \phi'(F \cap \bar{\phi}(T_\Gamma))$. En effet si $t \in \bar{\phi}^{-1}(F)$ alors $\bar{\phi}(t) \in F \cap \bar{\phi}(T_\Gamma)$ et $t = \phi'(\bar{\phi}(t)) \in \phi'(F \cap \bar{\phi}(T_\Gamma))$. Réciproquement, si $t \in \phi'(F \cap \bar{\phi}(T_\Gamma))$, c'est qu'il existe $t' \in T_\Gamma$ tel que $\bar{\phi}(t') \in F$ et $t = \phi'(\bar{\phi}(t')) = t'$, d'où $\bar{\phi}(t) = \bar{\phi}(t')$ est un élément de F et $t \in \bar{\phi}^{-1}(F)$.

Comme la classe des forêts algébriques est fermée par homomorphisme linéaire et par intersection avec une forêt reconnaissable (Rounds, 1970a), et que comme $\bar{\phi}$ est linéaire, $\bar{\phi}(T_\Gamma)$ est une forêt reconnaissable, $\bar{\phi}^{-1}(F) = \phi'(F \cap \bar{\phi}(T_\Gamma))$ est bien une forêt algébrique. C.Q.F.D.

Nous allons montrer maintenant que dans les hypothèses des théorèmes 4.1 et 4.2 la condition de linéarité est indispensable.

Considérons les deux alphabets gradués Σ et Δ avec $\Sigma_0 = \Delta_0 = \{\#\}$, $\Sigma_1 = \{a, b, c, d\}$, $\Delta_1 = \{a, b, c\}$, et $\Delta_2 = \{f\}$.

Soit le démarquage dyadique non linéaire ϕ de $T_\Sigma(X)$ dans $T_\Delta(X)$ défini par: $\phi(d(x)) = f(x, x)$; $\phi(a(x) = a(x)$; $\phi(b(x)) = b(x)$; $\phi(c(x)) = c(x)$; $\phi(\#) = \#$.

Soit la forêt algébrique $F = \{f(a^m b^m c^n \#, a^p b^p c^q \#) \mid m, n, p, q \geqslant 1\}$. Alors comme $\phi(da^{m_1} b^{m_2} c^{m_3} \#) = f(a^{m_1} b^{m_2} c^{m_3} \#, a^{m_1} b^{m_2} c^{m_3} \#)$, il est immédiat que $\phi^{-1}(F) = \{da^n b^n c^n \# \mid n \geqslant 1\}$ qui n'est manifestement pas algébrique.

5. Conclusions

On dira qu'un alphabet gradué Σ est n-adique si $\forall p \geqslant n$, $\Sigma_p = \varnothing$. On dira qu'un homomorphisme ϕ de $T_\Sigma(X)$ dans $T_\Delta(X)$ est n-adique si Σ et Δ sont des alphabets n-adiques. On appelle Alg_n la classe des forêts algébriques sur des alphabets n-adiques.

Il est évident que les classes Alg_n sont fermées par homomorphismes n-adiques linéaires, d'après Rounds (1970b) (mais pas par homomorphismes n-adiques non linéaires (Arnold et Dauchet, 1976; Engelfriet et Schmidt, 1975).

Les résultats obtenus dans cet article permettent d'ajouter les précisions suivantes

(1) Alg_1 est fermée par homomorphisme monadique linéaire inverse,

(2) Alg_2 est fermée par homomorphisme dyadique linéaire inverse,

(3) pour $n \geqslant 3$, Alg_n n'est pas fermée par homomorphisme n-adique linéaire inverse.

Comme la classes Alg_1 est pratiquement identique à celle des langages algébriques, le résultat (1) n'est cité que pour des raisons esthétiques, car il n'a absolument rien d'étonnant. Par contre la différence entre les résultats (2) et (3) amène à penser que la sous-classe Alg_2 de la classe des forêts algébriques doit posséder des propriétés particulières qui l'apparentent à celle des langages algébriques, et doit en conséquence être plus agréable à manipuler que la classe Alg toute entière.

Par ailleurs le résultat (3) nous a amené à considérer la plus petite famille de forêts contenant Alg et fermée par homomorphisme linéaire, homomorphisme linéaire inverse et intersection avec une forêt reconnaissable. Nous montrons (Arnold, 1977) que cette classe est une extension parfaitement naturelle de la classe Alg et que ses forêts peuvent être engendrées par des grammaires algébriques où un symbole non terminal peut être réécrit par un n-uple d'arbres.

RECEIVED: November 28, 1975; REVISED: July 7, 1977

REFERENCE

ARNOLD, A. (1977), "Systèmes d'équations dans le magmoïde. Ensembles rationnels et algébriques d'arbres," Thèse d'État, Lille.

ARNOLD, A., ET DAUCHET, M. (1975), Un théorème de Chomsky-Schützenberger pour les forêts algébriques, *Calcolo* **14**, 161–184.

ARNOLD, A., ET DAUCHET, M. (1976), Un théorème de duplication pour les forêts algébriques, *J. Comput. System Sci.* **13**, 223–244.

DAUCHET, M. (1975), "Transductions inversibles de forêts," Thèse 3° cycle, Lille.

ENGELFRIET, J. (1975), "Tree Automata and Tree Grammars," Daimi-Report FN-10, Univ. of Aarhus, Danemark.

ENGELFRIET, J., ET SCHMIDT, E. M. (1975), "IO and OI," Daimi-Report PB-47. Univ. of Aarhus, Danemark.

GINSBURG, S., ET GREIBACH, S. (1969), Abstract families of languages, *in* "Studies in Abstract Families of Languages," Memoir 87, pp. 1–32, Amer. Math. Soc., Providence, R.I.

MAIBAUM, T. S. E. (1974), A generalized approach to formal languages, *J. Comput. System Sci.* **8**, 409–439.

NIVAT, M. (1973), Langages algébriques sur le magma libre et sémantique des schémas de programme, *in* "Automata, Languages and Programming" (M. Nivat, Ed.), pp. 293–307, North-Holland, Amsterdam.

ROUNDS, W. C. (1969), Context-free grammars on trees, *in* "First Symp. on Theory of Computing," pp. 143–148.

ROUNDS, W. C. (1970a), Mappings and grammars on trees, *Math. Systems Theory* **4**, 257–287.

ROUNDS, W. C. (1970b), Tree oriented proofs of some theorems on context-free and indexed languages," *in* 2nd ACM Symp. on Theory of Computing," pp. 109–116.